



XIV COLOQUIO INTERNACIONAL SOBRE GESTIÓN UNIVERSITARIA

**La gestión del conocimiento
y los nuevos modelos de Universidades**

Florianópolis, 3 al 5 de Diciembre de 2014

CIGU2014

ÁREA TEMÁTICA 2 POLÍTICAS PÚBLICAS PARA EDUCACIÓN SUPERIOR

TÍTULO: FINANCIAMIENTO Y TAMAÑO DE LAS UNIVERSIDADES PÚBLICAS ARGENTINAS: UN ANÁLISIS DE CLUSTERING HARD Y DIFUSO

María Antonia Artola¹
Mariano Morettini²
Germán Blanco³

¹ martola@mdp.edu.ar, Universidad Nacional de Mar del Plata, Facultad de Ciencias Económicas y Sociales.

² mariano.morettini@gmail.com, Universidad Nacional de Mar del Plata, Facultad de Ciencias Económicas y Sociales.

³ white_german@hotmail.com, Universidad Nacional de Mar del Plata, Facultad de Ciencias Económicas y Sociales.

INDICE

Resumen.....	[03]
I.Introducción.....	[04]
II.Metodología utilizada.....	[04]
III.Resultados obtenidos.....	[06]
III.1.Clustering duro.....	[06]
III.2.Clustering difuso.....	[12]
IV.Conclusiones.....	[16]
V.Bibliografía consultada.....	[17]

RESUMEN

En base a los últimos datos públicamente disponibles en la Secretaría de Políticas Universitarias respecto de la cantidad de alumnos, docentes y recursos otorgados por el Tesoro Nacional a cada una de las universidades de gestión pública de la Argentina, nos proponemos aplicar técnicas alternativas de agrupamiento o *clustering* para caracterizar las universidades argentinas de gestión pública según su tamaño y participación en la distribución de recursos presupuestarios nacionales.

En particular, se utiliza el método tradicional de clustering “K-Medias” (*K-Means*) y se compara y enriquece con los resultados aportados por una técnica de agrupamiento borroso denominada “C-Medias Difuso” (“*Fuzzy C-Means*”). Esta última permite que cada institución pueda formar parte de cada agrupamiento en un grado de pertenencia parcial, en lugar de seguir la lógica booleana de simplemente pertenecer o no pertenecer.

El análisis pretende así conformar grupos de universidades homogéneos hacia adentro y heterogéneos entre sí, para luego, considerando el ámbito y la época de creación de las mismas, entre otras particularidades, realizar un aporte a la caracterización del sistema universitario público argentino.

Palabras clave: financiamiento universitario; clustering; fuzzy logic

I. Introducción

El objetivo del presente trabajo es agrupar a las distintas universidades nacionales argentinas teniendo en consideración su tamaño, para que los grupos así definidos puedan servir de base a futuras investigaciones relacionadas con la similitud entre instituciones, con el reparto del presupuesto público en el sistema de educación superior, con las distintas estrategias educativas perseguidas con la creación de universidades por parte de los gobiernos de turno o con cualquier otra temática que se vincule con la educación superior.

Para agrupar a las universidades, utilizamos como insumo la última información brindada por la Secretaría de Políticas Universitarias (SPU) y comparamos diferentes técnicas de agrupamiento. Por un lado, aplicamos una técnica de *clustering* tradicional o duro, y por otro lado, aplicamos una técnica de *clustering* difuso; con el objetivo de comparar y complementar los resultados y extraer conclusiones más enriquecedoras.

II. Metodología utilizada

En base a la última información disponible publicada por la SPU, referida al año 2011, hemos tomado los datos de la cantidad de cargos docentes, la cantidad de alumnos y el monto que el Tesoro Nacional le otorgó a cada una de las 42 Universidades Nacionales⁴ y hemos realizado algunos análisis de *clustering* con el objetivo de identificar grupos diferenciados de universidades.

Por una parte, la utilización de la técnica de *clustering* duro consiste en asignar a cada unidad de análisis (v.g. universidad), caracterizado por conjunto de atributos n dimensional (v.g. cantidad de alumnos, cantidad de docentes, monto transferido), a uno y solo un grupo o *cluster*; de manera de minimizar las diferencias n dimensionales de cada unidad de análisis respecto del elemento más representativo de dichos grupos, o centroide. Dichos *clusters* duros o tradicionales cumplen con dos condiciones básicas: abarcar todo el universo (esto es, que todas las unidades de análisis formen parte de un

⁴ De las 47 Universidades Nacionales que recibieron transferencias presupuestarias del Tesoro Nacional en el año 2011, hemos considerado solo 42 a los efectos del presente trabajo. No se han considerado las Universidades Nacionales Arturo Jauretche, de Avellaneda, de José C. Paz, del Oeste, de Tierra del Fuego y de Villa Mercedes, por no poseer, en el período de referencia docentes y/o alumnos, debido a su creación reciente y ausencia de funcionamiento educativo.

cluster) y que dichos *clusters* son excluyentes (esto es, que cada unidad de análisis forma parte de un solo *cluster*).

Así, en una primera instancia, utilizamos el método de *clustering* duro o tradicional denominado “K-medias” o “*K-means*”, que es del tipo supervisado. Esencialmente, esta técnica resulta en un proceso iterativo de asignación y reasignación de cada unidad de análisis a un centro de *cluster* o centroide, hasta minimizar la suma de alguna medida de distancia n dimensional entre cada unidad de análisis y los centros de *cluster*, que se calculan en cada iteración como el promedio de las distancias entre las dimensiones de cada unidad y las del centro de *cluster*.

Por su parte, las técnicas de *clustering* difuso se basan en similares criterios, pero por aplicación de las nociones de la lógica difusa, permiten que cada unidad de análisis sea asignada a cada grupo o *cluster* difuso, con un determinado grado de pertenencia; o alternativamente, que forme parte simultáneamente de más de un grupo difuso con un determinado grado de pertenencia.

Así, a fines de ampliar la información y el análisis, se aplicó el denominado “Fuzzy C-*means*”, tal como fue propuesto por Bezdek (1981) y que representa una versión de la técnica de *clustering* que utiliza lógica difusa.

Las variables seleccionadas para realizar el agrupamiento son las que interpretamos resultan más significativas para identificar grupos de universidades según su tamaño, esto es: cantidad de alumnos, cantidad de docentes y presupuesto recibido del Tesoro Nacional. Debido a que los valores absolutos de las tres variables son muy disímiles entre sí, previamente al procesamiento se procedió a estandarizar los valores⁵. La disparidad entre los valores absolutos de las variables, implica que el procedimiento iterativo elegido se vería fuertemente influenciado por la variable que mayores valores toma; es decir, el presupuesto asignado por el Tesoro Nacional, restándole importancia al resto de las variables, especialmente a la cantidad de docentes.

La elección del método de “K-means”, que es supervisado, se debe a que, debido al conocimiento sobre el tema que poseemos, entendemos que menos de 5 *clusters* resultaría inapropiado para una correcta diferenciación de algunas instituciones, mientras que más de 5 puede resultar en la formación de grupos no lo suficientemente

⁵ Para 2011 la media del presupuesto destinado por el Tesoro Nacional a las Universidades Nacionales fue de \$400,6 millones, la media de docentes fue de 3.870 y la media de estudiantes 33.406. La estandarización de las variables se realizó restando a cada dato la media respectiva y obteniendo el cociente con el respectivo desvío estándar.

heterogéneos entre sí. Más adelante haremos alguna justificación adicional acerca de la elección de 5 *clusters*.

Por otra parte, la aplicación del método difuso nos permite extraer algunas conclusiones adicionales. Como es sabido, la lógica difusa rompe con el principio aristotélico del tercero excluido, que indica que un elemento puede pertenecer o no pertenecer a un conjunto dado, vedándose cualquier otra posibilidad. Al trabajar con lógica difusa, se permite que un elemento pertenezca en forma parcial a un conjunto determinado, trabajándose así con grados de pertenencia de diferentes elementos a distintos conjuntos.

Ésta lógica nos permitiría concluir que una universidad pertenece al *cluster* de las universidades pequeñas, por ejemplo, en un grado 0,7 y al de universidades grandes en 0,3; lo que implicaría que, evidentemente, no es una universidad grande, pero seguramente no tan pequeña como otra cuyo grado de pertenencia a los mismos conjuntos sean 0,9 y 0,1, respectivamente. Como veremos luego, los resultados de aplicar estas técnicas enriquecen el análisis.

III. Resultados obtenidos

III.1. Clustering duro

Los resultados obtenidos de la aplicación de la técnica de “K-medias”, para $K=5$ y considerando como medida de la distancia a la euclidiana, utilizando el software SPSS®, se muestran en la Tabla 1, en la que puede observarse en qué *cluster* queda incluida cada una de las universidades.

En la Tabla 2 se propone un nombre para cada *cluster*, se muestra el promedio de las características n dimensionales de los *clusters*, se identifican las universidades que resultaron como los centros de cada *cluster* y la distancia (euclidiana) promedio de cada elemento de cada *cluster* a su respectivo centro. Por su parte, la distancia entre *clusters* se muestra en la Tabla 3.

Puede observarse que, obviando el primer *cluster* que está integrado solamente por la U.B.A, los *clusters* 3 y 4 presentan una significativa homogeneidad hacia adentro, a la vez que los *clusters* 2 y 5 muestran una menor homogeneidad hacia adentro, lo que puede verse en la última fila de la Tabla 2, donde mayores valores de distancia promedio representan menor homogeneidad de los elementos de cada *cluster*.

Tabla 1: Composición de *clusters*

<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>	<i>Cluster 4</i>	<i>Cluster 5</i>
UBA	Córdoba Rosario Tucumán	Centro PBA Comahue Cuyo La Matanza Litoral Lomas de Zamora Mar del Plata Misiones Nordeste Patagonia S.J.Bosco Río Cuarto Salta San Juan San Luis Sur	Catamarca Chaco Austral Chilecito Entre Ríos Formosa Gral. Sarmiento Inst.Univ.Arte Jujuy La Pampa La Rioja Lanús Luján Moreno Noroeste PBA Patagonia Austral Quilmes Río Negro San Martín Sgo.del Estero Tres de Febrero Villa María	La Plata UTN

Fuente: Elaboración propia en base a SPU

Tabla 2: Datos de los *clusters*

<i>Cluster</i>	1	2	3	4	5
Denominación	UBA	Grandes	Medianas	Pequeñas	Muy grandes
Centro <i>Cluster</i>	UBA	Rosario	San Juan	Moreno	UTN
Presup. Tesoro Prom.	\$ 3.076.026.243	\$ 978.228.995	\$ 361.091.771	\$ 138.858.899	\$ 1.241.240.832
Cargos Doc. Prom.	30.362	8.088	3.178	1.173	17.811
Estudiantes Prom.	351.200	80.693	25.912	10.941	95.675
Tamaño del <i>Cluster</i>	1	3	15	21	2
Dist. prom. al centro	0	0,5467	0,4697	0,3348	0,5680

Fuente: Elaboración propia

Tabla 3: Distancia entre *clusters*

Distancias entre		UBA	Grandes	Medianas	Pequeñas	Muy grandes
<i>Cluster</i>		1	2	3	4	5
UBA	1		7,411	9,153	9,750	6,178
Grandes	2	7,411		1,762	2,371	1,798
Medianas	3	9,153	1,762		0,614	3,316
Pequeñas	4	9,750	2,371	0,614		3,910
Muy grandes	5	6,178	1,798	3,316	3,910	

Fuente: Elaboración propia

Respecto a las distancias entre los centros de *clusters*, es una medida resumida de qué tan diferentes son los *clusters* entre sí, todos se alejan mucho del *cluster* 1 (v.g. la

UBA), según puede observarse en la primera columna de la Tabla 3. También puede observarse que entre el *cluster* 3 (universidades medianas) y el *cluster* 4 (universidades pequeñas) la distancia es relativamente baja. Por su parte, la distancia entre el *cluster* 2 (universidades grandes) y el *cluster* 5 (universidades muy grandes), también es relativamente baja, aunque es casi el triple que la anterior. Es de destacar, además, que aún en el caso de *clusters* cercanos, como los mencionados anteriormente, la distancia entre sus centros es siempre menor a la distancia promedio de cada elemento de un *cluster* a su respectivo centro.

Concluiríamos, entonces, con la información presentada en la Tabla 3, que las universidades pequeñas y medianas no son tan disímiles entre sí; y que tampoco lo son las grandes de las muy grandes. Sin embargo, sí existe fuerte heterogeneidad entre las pequeñas / medianas y las grandes / muy grandes.

Analizaremos ahora cada *cluster* para caracterizar adecuadamente a cada uno de los grupos formados:

- El primer *cluster* está compuesto sólo por la U.B.A. Posee 30.362 cargos docentes, cuando el promedio de las otras universidades nacionales es de 3.224; tiene una cantidad de alumnos equivalente al promedio de las demás instituciones multiplicado por 14 y dispone de un presupuesto que es más de un 800% superior al del promedio de todas las universidades nacionales. Puede pensarse que mientras que en tamaño (medido por docentes y alumnos) supera entre 10 y 14 veces a la universidad promedio, “solo” recibe 7 veces más presupuesto que la universidad promedio. La U.B.A es, así, la universidad más grande del país, la segunda en antigüedad, emplazada en la capital de un país que concentra al 7,2% de la población total en el 0,007% de la superficie del país, teniendo una densidad de población de 14.216 habitantes por km², siendo el promedio del país de 14,4 habitantes por km².

Debido a la particular condición del *cluster* 1, a continuación compararemos los valores de los diferentes *clusters* con el promedio de todas las instituciones, excluida la UBA.

- El segundo *cluster* puede ser denominado “Universidades Grandes”. Tiene como centro a la Universidad Nacional de Rosario, con una cantidad de cargos docentes y de estudiantes que duplican al promedio de las universidades, y con un presupuesto que representa el 260% del promedio. Las universidades que conforman este *cluster*, no sólo son grandes en función de la cantidad de docentes, de alumnos y de presupuesto, sino que la función social que

históricamente han desarrollado es significativa. Nótese que las tres se encuentran en enclaves geográficos estratégicos del país: Córdoba, a 700 km de la CABA; Rosario, a 300 km de la CABA y Tucumán, a cerca de 1200 km de la CABA. La ciudad de Córdoba tiene 1.300.000 habitantes, siendo el segundo centro urbano del país, Rosario tiene cerca de 1.000.000 y Tucumán 550.000. El año de creación de estas universidades ha sido 1613, 1968 y 1921, siendo la 1ª, 10ª y 5ª universidades nacionales creadas, respectivamente. Analizando la cantidad de población que habita las ciudades en que se asientan estas instituciones y la distancia a la capital del país, se explica que estén entre las 10 primeras universidades públicas creadas en el país, y que reciban de parte del Tesoro Nacional un presupuesto que sea más que proporcional a la cantidad de alumnos y de docentes que detentan.

- En el tercer *cluster*, que hemos denominado “Universidades Medianas”, encontramos 15 instituciones que en promedio tienen la misma cantidad de docentes y de estudiantes que el promedio de las universidades y un presupuesto del Tesoro un 7,6% superior al promedio. Encontramos aquí a muchas de las universidades creadas entre 1971-1980, período al que nos referiremos a continuación, a las Universidades del Sur (1956), del Nordeste (1956), del Litoral (1919) y la de La Matanza (1989).
- El cuarto *cluster* lo denominamos “Universidades Pequeñas”, y contiene a 21 instituciones, siendo su centroide la Universidad Nacional de Moreno. Estas universidades tienen, en promedio, un 64% menos de docentes y un 57% menos de estudiantes que el promedio y un 60% menos de presupuesto. Las universidades que componen este *cluster* pertenecen a diferentes procesos creadores que vivió el país, con un perfil particular cada uno de ellos. Por un lado se encuentran en este grupo todas las universidades creadas entre 1988 y 1995, con la excepción de La Matanza. Estas universidades se caracterizan por centralizarse mayormente en el área geográfica metropolitana, brindan una oferta especializada y no tradicionales, poseen una matrícula más limitada que las universidades nacionales anteriores y responden a un criterio más mercantilista, en el sentido de una tendencia hacia el autofinanciamiento, a través de cuotas voluntarias, carreras a distancia aranceladas, postgrados especializados pagos, convenios con municipalidades, actividades de transferencia al medio, entre otras alternativas (Pucheta y Morettini, 2010). Por

otra parte, también se encuentran todas las universidades creadas desde el año 2002 al presente, que tienen un perfil diferente al de las anteriormente mencionadas, pero que tienen una limitada zona de influencia y una escasísima antigüedad. Por último, encontramos en este *cluster* a la mayoría de las universidades creadas o nacionalizadas por los gobiernos militares, cuya función principal, según alguna doctrina especializada, ha sido la de descentralizar a las universidades tradicionales, restándoles fortaleza y facilitando el control de los estudiantes, en pos del objetivo de la seguridad nacional (Fernandez Lamarra, 2002).

- El quinto *cluster* posee sólo 2 universidades, y puede ser denominado “Universidades Muy Grandes”, siendo su centroide la UTN. Poseen más del quíntuple de docentes que el promedio, casi el cuádruple de alumnos y un presupuesto que es un 260% mayor al del promedio (en todos los casos, sin considerar a la U.B.A. Claramente, estas universidades muy grandes poseen mayor cantidad de docentes que de alumnos, proporcionalmente, y un presupuesto por docente que es menor al del promedio de las universidades. Las universidades que componen este grupo fueron creadas en 1905 (La Plata) y 1959 (UTN), siendo la 3ª y la 9ª universidad nacional creada en el país. La primera se encuentra en un centro geográfico muy importante, cuarto del país y nace como contrapunto de la UBA, con un perfil diferente: mientras la UBA seguía el modelo humboldtiano, orientado hacia la investigación científica pura, el de La Plata seguía una concepción anglosajona liberal, teniendo como objetivo la formación de docentes para educar a la población, de científicos que desarrollaran las industrias y de filósofos que condujeran al país (Buchbinder, 2010). La UTN, por su parte, nace con el objetivo de formar ingenieros a lo largo y ancho del país, teniendo Facultades Regionales en Buenos Aires, Córdoba, Mendoza, Rosario, Santa Fe, Bahía Blanca, La Plata, Tucumán y Avellaneda, para luego sumarse nuevos centros hasta la actualidad. La particularidad de centrar sus carreras en el ámbito de las ingenierías, la constituye en realidad en Instituto Universitario, pero debido a que originalmente fue creada con el nombre de Universidad Obrera Nacional (1948), se le permitió seguir conservando el nombre de Universidad cuando cambió su denominación por la actual en 1959. La UNLP justifica su presupuesto en la cantidad de alumnos y docentes que poseen, al estar en un centro geográfico de

alta densidad poblacional y poseer una historia académica de renombre. La UTN, por su parte, suma a estas características la particularidad de su presencia en diferentes conglomerados urbanos del país.

Ahora bien, ya hemos mencionado que la elección de la conformación de 5 *clusters* se debió a que nuestro conocimiento sobre el tema nos permite concluir que esa es la cantidad más conveniente de grupos a formar. Sin embargo, a continuación esgrimiremos algunas justificaciones cuantitativas para defender dicha postura, adicionales a la justificación cualitativa derivada de un conocimiento previo del tema.

Por un lado, nos interesa que el promedio de la distancia de cada punto al centroide de su *cluster* sea lo más bajo posible, porque de esa manera nos aseguramos que los *clusters* sean homogéneos hacia adentro. Por otro lado, también nos interesa que la distancia entre los centroides sea lo mayor posible, porque de esa manera obtenemos heterogeneidad entre los grupos formados. Así, en la Tabla 4 mostramos el promedio de las distancias de cada punto al centroide del *cluster* al que pertenece (v.g. distancias intra-*cluster*) y el promedio de las distancias entre centros de *clusters* (v.g. distancias inter-*clusters*), para distinta cantidad de número de *cluster*.

Tabla4: Distancias intra y entre *clusters*

K	Distancia intra- <i>cluster</i>	Distancia inter- <i>cluster</i>	Tasa Marginal
3	0,6853	6,3497	
4	0,5327	5,2617	0,1403
5	0,4012	4,6263	0,2070
6	0,3527	4,2180	0,1188
7	0,2688	3,8605	0,2347

Fuente: Elaboración propia

En dicha tabla puede observarse que a medida que se incrementa la cantidad de *clusters*, se reduce tanto la distancia intra-*cluster* como la distancia inter-*cluster*, lo que equivale a decir que los *clusters* se vuelven más homogéneos hacia adentro y menos heterogéneos entre ellos, por lo que mejora una de las condiciones buscadas, pero empeora la otra. Es por ello que calculamos una tasa marginal de intercambio a medida que crece la cantidad de *clusters*⁶, que mostraría conceptualmente cuanto “mejora” la

⁶ Como el cociente entre la diferencia de distancias intra- clusters dividido por la diferencia de distancias obtenida entre-clusters

homogeneidad (en términos de cambios de la distancia intra-*cluster*) por cada punto de “empeoramiento” de heterogeneidad (en términos de cambios en la distancia inter-*cluster*). Para distinta cantidad de agrupamientos, debiera optarse en principio por aquella que mostrara la mayor Tasa Marginal.

Como puede observarse en la Tabla 4 anterior, resulta que el agrupamiento de las universidades en 7 *clusters* es la que cumple la condición anterior, por poseer una tasa marginal mayor al resto (0,2347), siendo la segunda mejor alternativa para 5 *clusters*. Sin embargo, en términos prácticos la diferencia entre 5 y 7 *clusters* es la generación de numerosos *clusters* con un solo integrante. (v.g. U.B.A, Córdoba, La Plata, UTN), y 2 *clusters* con similar densidad de datos (las universidades medianas y las pequeñas), lo que no contribuye a nuestro análisis. Es por esta cuestión que nos hemos inclinado a considerar que 5 es la cantidad más adecuada de grupos a obtener en nuestro agrupamiento. Por otra parte, consideramos que la conformación de los *clusters* obtenida, con la descripción de cada uno de ellos, resulta coherente con las características propias de las universidades que conforman cada grupo.

III.2. Clustering difuso

Tal como se comentara oportunamente, nuestra intención es enriquecer los resultados obtenidos con las técnicas tradicionales complementándolas con técnicas derivadas del uso de la lógica difusa, para lo que se aplicó el método “Fuzzy C-Means”, para el agrupamiento en 5 *clusters* difusos, al igual que en el método tradicional, utilizando Matlab ®. Las coordenadas estandarizadas para los centroides obtenidos son los que se muestran en la Tabla 5.

Debido a que los valores presentes en la Tabla 5 corresponden a datos estandarizados, denominamos como “Universidades Medianas” al último *cluster* porque posee valores muy cercanos a 0 en todas sus variables, esto es, el centroide de este *cluster* es prácticamente la media del sistema educativo superior en cuanto a la cantidad de docentes, estudiantes y recursos recibidos del Tesoro Nacional.

Por su parte, al 3º lo denominamos “Universidades Pequeñas” por poseer valores negativos en todas sus variables y sensiblemente diferentes de 0, es decir, el centroide de este *cluster* posee sus variables con algo menos de medio desvío standard a la izquierda de la media.

Los *clusters* 2 y 4 poseen ambos centroides valores positivos de todas las variables, por lo que se trata de universidades más grandes que el promedio, es por ello que

denominamos a estos *clusters* “Universidades Grandes” y “Universidades Muy grandes”, distinguiendo entre ambos porque en el 4° los valores estandarizados de todas las variables superan a los del 2°.

Por último, al 1° *cluster* lo llamamos “UBA” porque el centroide de este *cluster* coincide con los valores correspondientes a dicha universidad y porque, como veremos, la UBA solo pertenece a este *cluster* y el grado de pertenencia del resto de las instituciones a este *cluster* es prácticamente nulo.

Tabla 5: Centroides Fuzzy C-Means

<i>Cluster</i>	Z Tesoro	Z Docentes	Z Estudiantes	Nombre <i>cluster</i>
1	5,1298	4,6451	5,6907	UBA
2	1,0607	0,8196	0,8204	Grandes
3	-0,4669	-0,4615	-0,3890	Pequeñas
4	1,6204	2,4894	1,0776	Muy grandes
5	-0,0120	-0,0256	-0,0820	Medianas

Fuente: Elaboración propia

Los grados de pertenencia de cada universidad a cada *cluster* difuso se muestran en la Tabla 6, en forma comparativa con los resultados del agrupamiento duro ya comentado. Allí podemos observar que cada universidad posee un grado de pertenencia en $[0;1]$ a cada *cluster*, donde 1 indica pertenencia absoluta y 0 no pertenencia absoluta. A efectos de hacer más representativa la pertenencia difusa, se ha optado por pintar cada grado de pertenencia asignándoles un tono más oscuro a las mayores pertenencias y más claras a las menores. Así, por ejemplo, no habría dudas de que la Universidad Nacional de Catamarca es pequeña, porque pertenece a dicho *cluster* en un grado de 0,9634, muy cercano a 1; pero no está tan claro, por ejemplo, el caso de la Universidad Nacional de La Matanza, que es casi tanto pequeña como mediana.

Podemos “deffuzyficar” los resultados, asignando cada universidad al *cluster* donde presenta mayor grado de pertenencia. Así, la Universidad de Salta sería Pequeña (porque tiene un grado de pertenencia mayor para ese *cluster* que para el resto; 0,5463) y pertenecería a dicho *cluster* tanto como la Universidad Nacional de Formosa, cuyo grado de pertenencia al mismo es 0,9960, pero evidentemente se estaría desaprovechando la información obtenida con esta técnica.

Como puede observarse en la Tabla 6, el agrupamiento borroso agrega información complementaria al agrupamiento clásico. Así por ejemplo, la Universidad Nacional del

Comahue es esencialmente Mediana (presenta un mayor grado de pertenencia a este *cluster* que a cualquier otro, en los que tiene un grado de pertenencia casi nulo). Por su parte, la de Cuyo presenta también mayor grado de pertenencia a dicho grupo de universidades Medianas; pero sin embargo muestra ciertas características que las acerca borrosamente, en menor medida, a las Grandes y a las Pequeñas.

Tabla 6: Clustering Difuso y Duro

	Clustering Difuso						Clustering Duro					
	UBA	Grandes	Medianas	Pequeñas	MuyGrandes	Suma	UBA	Grandes	Medianas	Pequeñas	MuyGrandes	Suma
UBA	1,00					1,00	1,00					1,00
Catamarca		0,00	0,03	0,96	0,01	1,00				1,00		1,00
Centro de la PBA		0,01	0,32	0,66	0,00	1,00			1,00			1,00
Chaco Austral		0,01	0,08	0,90	0,00	1,00				1,00		1,00
Chilecito		0,01	0,07	0,92	0,00	1,00				1,00		1,00
Comahue		0,00	0,98	0,02		1,00			1,00			1,00
Córdoba	0,01	0,68	0,07	0,04	0,20	1,00		1,00				1,00
Cuyo	0,00	0,17	0,65	0,15	0,03	1,00			1,00			1,00
Entre Ríos		0,01	0,18	0,81	0,00	1,00				1,00		1,00
Formosa			0,00	1,00		1,00				1,00		1,00
Gral. Sarmiento		0,01	0,04	0,95	0,00	1,00				1,00		1,00
Instituto Univ. Arte		0,01	0,20	0,78	0,00	1,00				1,00		1,00
Jujuy		0,00	0,02	0,97		1,00				1,00		1,00
La Matanza	0,00	0,03	0,48	0,48	0,01	1,00			1,00			1,00
La Pampa		0,00	0,05	0,94	0,00	1,00				1,00		1,00
La Plata	0,01	0,17	0,04	0,02	0,76	1,00					1,00	1,00
La Rioja		0,02	0,34	0,63	0,01	1,00				1,00		1,00
Lanús		0,00	0,02	0,97	0,00	1,00				1,00		1,00
Litoral	0,00	0,05	0,86	0,09	0,01	1,00			1,00			1,00
Lomas de Zamora	0,00	0,03	0,79	0,18	0,01	1,00			1,00			1,00
Luján		0,01	0,10	0,89	0,00	1,00				1,00		1,00
Mar del Plata		0,02	0,89	0,08	0,01	1,00			1,00			1,00
Misiones		0,01	0,29	0,69	0,00	1,00			1,00			1,00
Moreno	0,00	0,02	0,10	0,87	0,01	1,00				1,00		1,00
Nordeste	0,00	0,10	0,76	0,12	0,02	1,00			1,00			1,00
Noroeste de la PBA		0,01	0,04	0,95	0,00	1,00				1,00		1,00
Patagonia Austral			0,01	0,99		1,00				1,00		1,00
Patagonia S. J. Bosco	0,00	0,02	0,67	0,31	0,01	1,00			1,00			1,00
Quilmes		0,00	0,02	0,97		1,00				1,00		1,00
Río Cuarto		0,01	0,30	0,69	0,00	1,00			1,00			1,00
Río Negro		0,01	0,05	0,94	0,00	1,00				1,00		1,00
Rosario	0,00	0,96	0,02	0,01	0,01	1,00		1,00				1,00
Salta	0,00	0,02	0,43	0,55	0,01	1,00			1,00			1,00
San Juan	0,00	0,03	0,84	0,12	0,01	1,00			1,00			1,00
San Luis		0,01	0,23	0,75	0,00	1,00			1,00			1,00
San Martín		0,00	0,04	0,96		1,00				1,00		1,00
Sgo del Estero		0,00	0,01	0,99		1,00				1,00		1,00
Sur		0,01	0,79	0,19	0,00	1,00			1,00			1,00
UTN	0,01	0,05	0,02	0,01	0,91	1,00					1,00	1,00
Tres de Febrero		0,00	0,02	0,97	0,00	1,00				1,00		1,00
Tucumán	0,00	0,59	0,25	0,10	0,05	1,00		1,00				1,00
Villa María		0,01	0,04	0,95	0,00	1,00				1,00		1,00

Fuente: Elaboración propia

La riqueza del método consiste en sacarle provecho a esta información gradual que se obtiene. Podría hacerse, por ejemplo, un ranking de las distintas universidades conforme su grado de pertenencia a cada *cluster*. Así, el *cluster* de las Universidades Medianas estaría encabezado por la Universidad Nacional del Comahue (es la “más mediana” de todas, porque es la que mayor grado de pertenencia a ese *cluster* presente, 0,98), seguida por la de Mar del Plata (0,89), la del Litoral (0,86), la de San Juan (0,84), y así sucesivamente. La universidad que menor grado de pertenencia tiene a este *cluster*, y sin embargo debe considerarse mediana por tener incluso menores grados de pertenencia a los otros *clusters*, es la de Cuyo (0,65). Similares análisis pueden hacerse con respecto a los otros *clusters*.

Por su parte, la Tabla 7 muestra el *cluster* al que pertenece cada Universidad de acuerdo a la metodología dura y a la difusa (para el caso de asignarse cada universidad al *cluster* en el que muestra el mayor grado de pertenencia. Nótese que solo 6 instituciones cambian de *cluster*, y todas para pasar de Medianas (en el *clustering* duro) a Pequeñas (en el *clustering* difuso): Centro de la Provincia de Buenos Aires, La Matanza, Misiones, Río Cuarto, Salta y San Luis.

Tabla 7: Comparación de pertenencia a *clusters* con metodología dura y difusa

Universidad	Clustering duro vs. difuso		
	Duro	Difuso	Cambio
Buenos Aires	UBA	UBA	
Catamarca	Pequeña	Pequeña	
Centro de la PBA	Mediana	Pequeña	Si
Chaco Austral	Pequeña	Pequeña	
Chilecito	Pequeña	Pequeña	
Comahue	Mediana	Mediana	
Córdoba	Grande	Grande	
Cuyo	Mediana	Mediana	
Entre Ríos	Pequeña	Pequeña	
Formosa	Pequeña	Pequeña	
Gral. Sarmiento	Pequeña	Pequeña	
Instituto Univ. del Arte	Pequeña	Pequeña	
Jujuy	Pequeña	Pequeña	
La Matanza	Mediana	Pequeña	Si
La Pampa	Pequeña	Pequeña	
La Plata	Muy grande	Muy grande	
La Rioja	Pequeña	Pequeña	
Lanús	Pequeña	Pequeña	
Litoral	Mediana	Mediana	
Lomas de Zamora	Mediana	Mediana	
Luján	Pequeña	Pequeña	

Universidad	Clustering duro vs. difuso		
	Duro	Difuso	Cambio
Mar del Plata	Mediana	Mediana	
Misiones	Mediana	Pequeña	Si
Moreno	Pequeña	Pequeña	
Nordeste	Mediana	Mediana	
Noroeste de la PBA	Pequeña	Pequeña	
Patagonia Austral	Pequeña	Pequeña	
Patagonia S. J. Bosco	Mediana	Mediana	
Quilmes	Pequeña	Pequeña	
Río Cuarto	Mediana	Pequeña	Si
Río Negro	Pequeña	Pequeña	
Rosario	Grande	Grande	
Salta	Mediana	Pequeña	Si
San Juan	Mediana	Mediana	
San Luis	Mediana	Pequeña	Si
San Martín	Pequeña	Pequeña	
Santiago del Estero	Pequeña	Pequeña	
Sur	Mediana	Mediana	
Tecnológica Nacional	Muy grande	Muy grande	
Tres de Febrero	Pequeña	Pequeña	
Tucumán	Grande	Grande	
Villa María	Pequeña	Pequeña	

Fuente: Elaboración propia

IV. Conclusiones

El objetivo perseguido por el presente trabajo puede decirse que era doble: uno de resultado y el otro metodológico. El objetivo de resultado era obtener un agrupamiento de las universidades nacionales según su tamaño y participación en el financiamiento del Tesoro Nacional, para lo cual se consideró la última información disponible brindada por la Secretaría de Políticas Universitarias, para el año 2011, y se reputaron determinantes las variables: cantidad de alumnos, cantidad de docentes y monto recibido del Tesoro Nacional, para cada institución.

El objetivo metodológico consistía en comparar y complementar los resultados obtenidos por dos técnicas de *clustering* supervisado: el tradicional *Hard K-Means* y el *Fuzzy C-Means*, que es una variante del anterior donde se utiliza lógica difusa.

Las conclusiones a las que podemos arribar son múltiples. Podemos decir, por un lado, que se satisfizo el objetivo de agrupar a las universidades según su tamaño. Se hicieron algunas consideraciones metodológicas respecto de la importancia de estandarizar los datos al haber variables con valores absolutos muy diferentes entre sí y se comparó la utilización de técnicas tradicionales y difusas sin que hubiere demasiadas variaciones respecto de la composición de cada *cluster*, lo cual robustece los resultados obtenidos.

Sin embargo, consideramos que la principal conclusión que debemos extraer radica en la interpretación complementaria de los resultados del *clustering* difuso o borroso. La riqueza del método difuso se encuentra en interpretar adicionalmente los grados de pertenencia que las instituciones presentan a los distintos *clusters* formados, de manera tal que existen casos de universidades que muestran características propias de más de un grupo.

Para finalizar, diremos que para abordar un tema de análisis, la mejor y mayor información puede obtenerse a partir del tratamiento del mismo con diferentes técnicas que aporten alguna conclusión adicional para el mejor entendimiento del problema. Es por eso que debe hacerse un esfuerzo por trabajar con una concepción integradora y holística de análisis, planteando las distintas posibilidades de complementariedad de técnicas antes que su rivalidad.

El presente trabajo intentó ser un ejemplo de ello y, al mismo tiempo, brindar información acerca del tamaño de las universidades, que puede ser insumo para futuros trabajos tanto de los mismos autores como de otros investigadores.

V. Bibliografía Consultada

- Bezdek, James C. (1981): *Pattern recognition with fuzzy objective function algorithm*. Plenum Press , New York.
- Bezdek, James C. (1984): “FCM: The fuzzy c-means clustering algorithm”. In *Computers & Geosciences*, Vol. 10, Nº 2-3, pp. 191-203.
- Buchbinder, Pablo (2010): *Historia de las Universidades Argentinas*. 2º edición. Sudamericana, Buenos Aires.
- Johnson, Dallas E. (2004): *Métodos multivariados aplicados al análisis de datos*. Soluciones empresariales, México.
- Fernandez Lamarra, Norberto (2002): “La educación superior en la Argentina”. Buenos Aires.
- Pal, Nikhil R and Bezdek, James C. (1995): “On cluster validity for the fuzzy c-means model”. In *IEEE Transactions on Fuzzy Systems*, Vol. 3, Nº 3, pp. 370-379.
- Pucheta, María Celeste y Morettini, Mariano (2010): “Aproximación al contexto histórico del proceso creador de universidades en el período 1989-1995 en Argentina”. En *Anales del X Congreso Internacional sobre Gestión Universitaria de América del Sur*. Mar del Plata.
- Secretaría de Políticas Universitarias (2011): “Anuario de Estadísticas Universitarias – Argentina 2011”. Ministerio de Educación de la Nación. Buenos Aires.
- Uriel, Ezequiel y Aldás, Joaquín (2005): *Análisis multivariante aplicado*. Thomson, Madrid.
- Yang, M.S. (1993): “A survey of fuzzy clustering”. In *Mathematical and Computer Modelling*, Vol. 18, Nº 11, pp. 1-16.
- Zadeh, Lofti A. (1965): “Fuzzy sets”. In *Information and Control*, Nº 8, pp. 338-353.